

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Mosselman, Sieste Dijkema, Rein
- (ii) TITLE OF INVENTION: Novel estrogen receptor
- (iii) NUMBER OF SEQUENCES: 28
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Akzo Nobel Patent Dept.
 - (B) STREET: 1300 Piccard Drive, Suite 206
 - (C) CITY: Rockville
 - (D) STATE: Maryland
 - (E) COUNTRY: US
 - (F) ZIP: 20850
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 08/826,361
 - (B) FILING DATE: 26-MAR-1997
 - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Gormley, Mary E.
 - (B) REGISTRATION NUMBER: 34,409
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 301-948-7400
 - (B) TELEFAX: 301-948-9751

(2) INFORMATION FOR SEQ ID NO: 1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1434 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

ATGAATTACA	GCATTCCCAG	CAATGTCACT	AACTTGGAAAG	GTGGGCCTGG	TCGGCAGACC	60
ACAAGCCCAA	ATGTGTTGTG	GCCAACACCT	GGGCACCTT	CTCCTTAGT	GGTCATCGC	120
CAGTTATCAC	ATCTGTATGC	GAACCTCAA	AAGAGTCCCT	GGTGTGAAGC	AAGATCGCTA	180
GAACACACCT	TACCTGTAAA	CAGAGAGACA	CTGAAAAGGA	AGGTTAGTGG	GAACCGTTGC	240
GCCAGCCCTG	TTACTGGTCC	AGGTTCAAAG	AGGGATGCTC	ACTTCTGCGC	TGTCTGCAGC	300
GATTACGCAT	CGGGATATCA	CTATGGAGTC	TGGTCGTGTG	AAGGATGTAA	GGCCTTTTTT	360
AAAAGAACCA	TTCAAGGACA	TAATGATTAT	ATTGTCCAG	CTACAAATCA	GTGTACAATC	420
GATAAAAACC	GGCGCAAGAG	CTGCCAGGCC	TGCCGACTTC	GGAAGTGTAA	CGAAAGTGGGA	480
ATGGTGAAGT	GTGGCTCCCG	GAGAGAGAGA	TGTGGGTACC	GCCTTGTGCG	GAGACAGAGA	540
AGTGCCGACG	AGCAGCTGCA	CTGTGCCGGC	AAGGCCAAGA	GAAGTGGCGG	CCACGCGCCC	600
CGAGTGCAGG	AGCTGCTGCT	GGACGCCCTG	AGCCCCGAGC	AGCTAGTGCT	CACCCTCCTG	660
GAGGCTGAGC	CGCCCCATGT	GCTGATCAGC	CGCCCGAGTG	CGCCCTTCAC	CGAGGCCTCC	720
ATGATGATGT	CCCTGACCAA	GTTGGCCGAC	AAGGAGTTGG	TACACATGAT	CAGCTGGGCC	780
AAGAAAGATT	CCGGCTTTGT	GGAGCTCAGC	CTGTTGACCC	AAGTGCAGCT	CTTGGAGAGC	840
TGTTGGATGG	AGGTGTTAAT	GATGGGGCTG	ATGTGGCGCT	CAATTGACCA	CCCCGGCAAG	900
CTCATTTTG	CTCCAGATCT	TGTTCTGGAC	AGGGATGAGG	GGAAATGCGT	AGAAGGAATT	960
CTGGAAATCT	TTGACATGCT	CCTGGCAACT	ACTTCAAGGT	TTCGAGAGTT	AAAACCTCAA	1020
CACAAAGAAT	ATCTCTGTGT	CAAGGCCATG	ATCCTGCTCA	ATTCAGTAT	GTACCCTCTG	1080
GTCACAGCGA	CCCAGGATGC	TGACAGCAGC	CGGAAGCTGG	CTCACTTGCT	GAACGCCGTG	1140
ACCGATGCTT	TGGTTGGGT	GATTGCCAAG	AGCGGCATCT	CCTCCAGCA	GCAATCCATG	1200
CGCCTGGCTA	ACCTCCTGAT	GCTCCTGTCC	CACGTCAAGC	ATGCGACTAA	CAAGGGCATG	1260
GAACATCTGC	TCAACATGAA	GTGAAAAAT	GTGGTCCCAG	TGTATGACCT	GCTGCTGGAG	1320
ATGCTGAATG	CCCACGTGCT	TCGCGGGTGC	AAGTCCTCCA	TCACGGGGTC	CGAGTGCAGC	1380
CGGGCAGAGG	ACAGTAAAAG	CAAAGAGGGC	TCCCAGAAC	CACAGTCTCA	GTGA	1434

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1251 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

ATGAATTACA	GCATTCCAG	CAATGTCACT	AACTTGGAAAG	GTGGGCCTGG	TCGGCAGACC	60
ACAAGCCCAA	ATGTGTTGTG	GCCAACACCT	GGGCACCTTT	CTCCTTAGT	GGTCCATCGC	120
CAGTTATCAC	ATCTGTATGC	GGAACCTCAA	AAGAGTCCCT	GGTGTGAAGC	AAGATCGCTA	180
GAACACACCT	TACCTGTAAA	CAGAGAGACA	CTGAAAAGGA	AGGTTAGTGG	GAACCGTTGC	240
GCCAGCCCTG	TTACTGGTCC	AGGTTCAAAG	AGGGATGCTC	ACTTCTGCGC	TGTCTGCAGC	300
GATTACGCAT	CGGGATATCA	CTATGGAGTC	TGGTCGTGTG	AAGGATGTAA	GGCCTTTTT	360
AAAAGAACCA	TTCAAGGACA	TAATGATTAT	ATTGTCCAG	CTACAAATCA	GTGTACAATC	420
GATAAAAACC	GGCGCAAGAG	CTGCCAGGCC	TGCCGACTTC	GGAAGTGTAA	CGAAGTGGGA	480
ATGGTGAAGT	GTGGCTCCCG	GAGAGAGAGA	TGTGGGTACC	GCCTTGTGCG	GAGACAGAGA	540
AGTGCCGACG	AGCAGCTGCA	CTGTGCCGGC	AAGGCCAAGA	GAAGTGGCGG	CCACCGGCC	600
CGAGTGCAGG	AGCTGCTGCT	GGACGCCCTG	AGCCCCGAGC	AGCTAGTGCT	CACCCTCCTG	660
:GAGGCTGAGC	CGCCCCATGT	GCTGATCAGC	CGCCCCAGTG	CGCCCTTCAC	CGAGGCCTCC	720
ATGATGATGT	CCCTGACCAA	GTTGGCCGAC	AAGGAGTTGG	TACACATGAT	CAGCTGGGCC	780
AAGAAGATTC	CCGGCTTGT	GGAGCTCAGC	CTGTTCGACC	AAGTGCAGCT	CTTGGAGAGC	840
TGTTGGATGG	AGGTGTTAAT	GATGGGGCTG	ATGTGGCGCT	CAATTGACCA	CCCCGGCAAG	900
CTCATCTTG	CTCCAGATCT	TGTTCTGGAC	AGGGATGAGG	GGAAATGCGT	AGAAGGAATT	960
CTGGAAATCT	TTGACATGCT	CCTGGCAACT	ACTTCAAGGT	TTCGAGAGTT	AAAACCTCAA	1020
SACAAAGAAT	ATCTCTGTGT	CAAGGCCATG	ATCCTGCTCA	ATTCCAGTAT	GTACCCTCTG	1080
GTCACAGCGA	CCCAGGATGC	TGACAGCAGC	CGGAAGCTGG	CTCACTTGCT	GAACGCCGTG	1140

ACCGATGCTT TGGTTGGGT GATTGCCAAG AGCGGCATCT CCTCCCAGCA GCAATCCATG 1200
CGCCTGGCTA ACCTCCTGAT GCTCCTGTCC CACGTCAGGC ATGCGAGGTG A 1251

(2) INFORMATION FOR SEQ ID NO: 3:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 66 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Cys Ala Val Cys Ser Asp Tyr Ala Ser Gly Tyr His Tyr Gly Val Trp
1 5 10 15

Ser Cys Glu Gly Cys Lys Ala Phe Phe Lys Arg Ser Ile Gln Gly His
20 25 30

Asn Asp Tyr Ile Cys Pro Ala Thr Asn Gln Cys Thr Ile Asp Lys Asn
35 40 45

Arg Arg Lys Ser Cys Gln Ala Cys Arg Leu Arg Lys Cys Tyr Glu Val
50 55 60

Gly Met
65

(2) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 233 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

Leu Val Leu Thr Leu Leu Glu Ala Glu Pro Pro His Val Leu Ile Ser
1 5 10 15

Arg Pro Ser Ala Pro Phe Thr Glu Ala Ser Met Met Met Ser Leu Thr
20 25 30

Lys Leu Ala Asp Lys Glu Leu Val His Met Ile Ser Trp Ala Lys Lys
35 40 45

Ile Pro Gly Phe Val Glu Leu Ser Leu Phe Asp Gln Val Arg Leu Leu
50 55 60

Glu Ser Cys Trp Met Glu Val Leu Met Met Gly Leu Met Trp Arg Ser
65 70 75 80

Ile Asp His Pro Gly Lys Leu Ile Phe Ala Pro Asp Leu Val Leu Asp
85 90 95

Arg Asp Glu Gly Lys Cys Val Glu Gly Ile Leu Glu Ile Phe Asp Met
100 105 110

Leu Leu Ala Thr Thr Ser Arg Phe Arg Glu Leu Lys Leu Gln His Lys
115 120 125

Glu Tyr Leu Cys Val Lys Ala Met Ile Leu Leu Asn Ser Ser Met Tyr
130 135 140

Pro Leu Val Thr Ala Thr Gln Asp Ala Asp Ser Ser Arg Lys Leu Ala
145 150 155 160

His Leu Leu Asn Ala Val Thr Asp Ala Leu Val Trp Val Ile Ala Lys
165 170 175

Ser Gly Ile Ser Ser Gln Gln Ser Met Arg Leu Ala Asn Leu Leu
180 185 190

Met Leu Leu Ser His Val Arg His Ala Ser Asn Lys Gly Met Glu His
195 200 205

Leu Leu Asn Met Lys Cys Lys Asn Val Val Pro Val Tyr Asp Leu Leu
210 215 220

Leu Glu Met Leu Asn Ala His Val Leu
225 230

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 477 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

Met	Asn	Tyr	Ser	Ile	Pro	Ser	Asn	Val	Thr	Asn	Leu	Glu	Gly	Gly	Pro
1				5					10						15
Gly	Arg	Gln	Thr	Thr	Ser	Pro	Asn	Val	Leu	Trp	Pro	Thr	Pro	Gly	His
				20					25						30
Leu	Ser	Pro	Leu	Val	Val	His	Arg	Gln	Leu	Ser	His	Leu	Tyr	Ala	Glu
				35				40							45
Pro	Gln	Lys	Ser	Pro	Trp	Cys	Glu	Ala	Arg	Ser	Leu	Glu	His	Thr	Leu
				50		55					60				
Pro	Val	Asn	Arg	Glu	Thr	Leu	Lys	Arg	Lys	Val	Ser	Gly	Asn	Arg	Cys
	65				70					75					80
Ala	Ser	Pro	Val	Thr	Gly	Pro	Gly	Ser	Lys	Arg	Asp	Ala	His	Phe	Cys
				85				90							95
Ala	Val	Cys	Ser	Asp	Tyr	Ala	Ser	Gly	Tyr	His	Tyr	Gly	Val	Trp	Ser
	100							105					110		
Cys	Glu	Gly	Cys	Lys	Ala	Phe	Phe	Lys	Arg	Ser	Ile	Gln	Gly	His	Asn
	115					120						125			
Asp	Tyr	Ile	Cys	Pro	Ala	Thr	Asn	Gln	Cys	Thr	Ile	Asp	Lys	Asn	Arg
	130					135					140				
Arg	Lys	Ser	Cys	Gln	Ala	Cys	Arg	Leu	Arg	Lys	Cys	Tyr	Glu	Val	Gly
	145				150					155					160
Met	Val	Lys	Cys	Gly	Ser	Arg	Arg	Glu	Arg	Cys	Gly	Tyr	Arg	Leu	Val
				165					170						175
Arg	Arg	Gln	Arg	Ser	Ala	Asp	Glu	Gln	Leu	His	Cys	Ala	Gly	Lys	Ala
		180				185							190		
Lys	Arg	Ser	Gly	Gly	His	Ala	Pro	Arg	Val	Arg	Glu	Leu	Leu	Leu	Asp
		195					200					205			

Ala Leu Ser Pro Glu Gln Leu Val Leu Thr Leu Leu Glu Ala Glu Pro
210 215 220

Pro His Val Leu Ile Ser Arg Pro Ser Ala Pro Phe Thr Glu Ala Ser
225 230 235 240

Met Met Met Ser Leu Thr Lys Leu Ala Asp Lys Glu Leu Val His Met
245 250 255

Ile Ser Trp Ala Lys Lys Ile Pro Gly Phe Val Glu Leu Ser Leu Phe
260 265 270

Asp Gln Val Arg Leu Leu Glu Ser Cys Trp Met Glu Val Leu Met Met
275 280 285

Gly Leu Met Trp Arg Ser Ile Asp His Pro Gly Lys Leu Ile Phe Ala
290 295 300

Pro Asp Leu Val Leu Asp Arg Asp Glu Gly Lys Cys Val Glu Gly Ile
305 310 315 320

Leu Glu Ile Phe Asp Met Leu Leu Ala Thr Thr Ser Arg Phe Arg Glu
325 330 335

Leu Lys Leu Gln His Lys Glu Tyr Leu Cys Val Lys Ala Met Ile Leu
340 345 350

Leu Asn Ser Ser Met Tyr Pro Leu Val Thr Ala Thr Gln Asp Ala Asp
355 360 365

Ser Ser Arg Lys Leu Ala His Leu Leu Asn Ala Val Thr Asp Ala Leu
370 375 380

Val Trp Val Ile Ala Lys Ser Gly Ile Ser Ser Gln Gln Gln Ser Met
385 390 395 400

Arg Leu Ala Asn Leu Leu Met Leu Leu Ser His Val Arg His Ala Ser
405 410 415

Asn Lys Gly Met Glu His Leu Leu Asn Met Lys Cys Lys Asn Val Val
420 425 430

Pro Val Tyr Asp Leu Leu Leu Glu Met Leu Asn Ala His Val Leu Arg
435 440 445

Gly Cys Lys Ser Ser Ile Thr Gly Ser Glu Cys Ser Pro Ala Glu Asp
450 455 460

Ser Lys Ser Lys Glu Gly Ser Gln Asn Pro Gln Ser Gln
465 470 475

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 416 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Met	Asn	Tyr	Ser	Ile	Pro	Ser	Asn	Val	Thr	Asn	Leu	Glu	Gly	Gly	Pro
1									10						15
Gly	Arg	Gln	Thr	Thr	Ser	Pro	Asn	Val	Leu	Trp	Pro	Thr	Pro	Gly	His
									25						30
Leu	Ser	Pro	Leu	Val	Val	His	Arg	Gln	Leu	Ser	His	Leu	Tyr	Ala	Glu
							35		40					45	
Pro	Gln	Lys	Ser	Pro	Trp	Cys	Glu	Ala	Arg	Ser	Leu	Glu	His	Thr	Leu
						50		55			60				
Pro	Val	Asn	Arg	Glu	Thr	Leu	Lys	Arg	Lys	Val	Ser	Gly	Asn	Arg	Cys
						65		70		75					80
Ala	Ser	Pro	Val	Thr	Gly	Pro	Gly	Ser	Lys	Arg	Asp	Ala	His	Phe	Cys
						85			90					95	
Ala	Val	Cys	Ser	Asp	Tyr	Ala	Ser	Gly	Tyr	His	Tyr	Gly	Val	Trp	Ser
						100			105				110		
Cys	Glu	Gly	Cys	Lys	Ala	Phe	Phe	Lys	Arg	Ser	Ile	Gln	Gly	His	Asn
							115		120			125			
Asp	Tyr	Ile	Cys	Pro	Ala	Thr	Asn	Gln	Cys	Thr	Ile	Asp	Lys	Asn	Arg
						130		135			140				
Arg	Lys	Ser	Cys	Gln	Ala	Cys	Arg	Leu	Arg	Lys	Cys	Tyr	Glu	Val	Gly
						145		150		155					160
Met	Val	Lys	Cys	Gly	Ser	Arg	Arg	Glu	Arg	Cys	Gly	Tyr	Arg	Leu	Val
						165			170					175	
Arg	Arg	Gln	Arg	Ser	Ala	Asp	Glu	Gln	Leu	His	Cys	Ala	Gly	Lys	Ala
						180			185				190		
Lys	Arg	Ser	Gly	Gly	His	Ala	Pro	Arg	Val	Arg	Glu	Leu	Leu	Leu	Asp
						195		200				205			

Ala Leu Ser Pro Glu Gln Leu Val Leu Thr Leu Leu Glu Ala Glu Pro
 210 215 220

Pro His Val Leu Ile Ser Arg Pro Ser Ala Pro Phe Thr Glu Ala Ser
 225 230 235 240

Met Met Met Ser Leu Thr Lys Leu Ala Asp Lys Glu Leu Val His Met
 245 250 255

Ile Ser Trp Ala Lys Lys Ile Pro Gly Phe Val Glu Leu Ser Leu Phe
 260 265 270

Asp Gln Val Arg Leu Leu Glu Ser Cys Trp Met Glu Val Leu Met Met
 275 280 285

Gly Leu Met Trp Arg Ser Ile Asp His Pro Gly Lys Leu Ile Phe Ala
 290 295 300

Pro Asp Leu Val Leu Asp Arg Asp Glu Gly Lys Cys Val Glu Gly Ile
 305 310 315 320

Leu Glu Ile Phe Asp Met Leu Leu Ala Thr Thr Ser Arg Phe Arg Glu
 325 330 335

Leu Lys Leu Gln His Lys Glu Tyr Leu Cys Val Lys Ala Met Ile Leu
 340 345 350

Leu Asn Ser Ser Met Tyr Pro Leu Val Thr Ala Thr Gln Asp Ala Asp
 355 360 365

Ser Ser Arg Lys Leu Ala His Leu Leu Asn Ala Val Thr Asp Ala Leu
 370 375 380

Val Trp Val Ile Ala Lys Ser Gly Ile Ser Ser Gln Gln Gln Ser Met
 385 390 395 400

Arg Leu Ala Asn Leu Leu Met Leu Leu Ser His Val Arg His Ala Arg
 405 410 415

(2) INFORMATION FOR SEQ ID NO: 7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 29 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: both
 - (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GGIGAYGARG CWTCIGGITG YCAYTAYGG

29

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

AAGCCTGGSA YICKYTTIGC CCAIYTIAT

29

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

TGTTACGAAG TGGGAATGGT GA

22

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

TTGACACCAG ACCAACTGGT AATG

24

(2) INFORMATION FOR SEQ ID NO: 11:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

GGTGGCGACG ACTCCTGGAG CCCG

24

(2) INFORMATION FOR SEQ ID NO: 12:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

GTACACTGAT TTGTAGCTGG AC

22

(2) INFORMATION FOR SEQ ID NO: 13:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

CCATGATGAT GTCCCTGACC

20

(2) INFORMATION FOR SEQ ID NO: 14:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

TAGCATGCCT GACGTGGGAC

20

(2) INFORMATION FOR SEQ ID NO: 15:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

GCCSTCCAGC ATCTCCAGSA RCAG

24

(2) INFORMATION FOR SEQ ID NO: 16:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 20 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

GGAAGCTGGC TCACTTGCTG

20

(2) INFORMATION FOR SEQ ID NO: 17:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

TCTTGTTC TG GACAGGGATG

20

(2) INFORMATION FOR SEQ ID NO: 18:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

GCATGGAACA TCTGCTAAC

20

(2) INFORMATION FOR SEQ ID NO: 19:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

AGCAAGTTCA GCCTGTTAAG T

21

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1257 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

ATGAATTACA	GCATTCCAG	CAATGTCACT	AACTTCCAAG	GTGGGCCTGG	TCGGCAGACC	60
ACAAGCCAA	ATGTGTTGTG	GCCAACACCT	GGGCACCTTT	CTCCTTAGT	GGTCCATCGC	120
CAGTTATCAC	ATCTGTATGC	GGAACCTCAA	AAGAGTCCCT	GGTGTGAAGC	AAGATCGCTA	180
GAACACACCT	TACCTGTAAA	CAGAGAGACA	CTGAAAAGGA	AGGTTAGTGG	GAACCGTTGC	240
GCCAGCCCTG	TTACTGGTCC	AGGTTCAAAG	AGGGATGCTC	ACTTCTGCGC	TGTCTGCAGC	300
GATTACGCAT	CGGGATATCA	CTATGGAGTC	TGGTCGTGTG	AAGGATGTAA	GGCCTTTTT	360
AAAAGAAGCA	TTCAAGGACA	TAATGATTAT	ATTTGTCCAG	CTACAAATCA	GTGTACAATC	420
GATAAAAACC	GGCGCAAGAG	CTGCCAGGCC	TGCCGACTTC	GGAAAGTGTAA	CGAAGTGGGA	480
ATGGTGAAGT	GTGGCTCCCG	GAGAGAGAGA	TGTGGGTACC	GCCTTGTGCG	GAGACAGAGA	540
AGTGCCGACG	AGCAGCTGCA	CTGTGCCGGC	AAGGCCAAGA	GAAGTGGCGG	CCACGCGCCC	600
CGAGTGCAGG	AGCTGCTGCT	GGACGCCCTG	AGCCCCGAGC	AGCTAGTGCT	CACCCCTCTG	660
GAGGCTGAGC	CGCCCCATGT	GCTGATCAGC	CGCCCCAGTG	CGCCCTTCAC	CGAGGCCTCC	720
ATGATGATGT	CCCTGACCAA	GTTGGCCGAC	AAGGAGTTGG	TACACATGAT	CAGCTGGGCC	780
AAGAAAGATT	CCGGCTTGT	GGAGCTCAGC	CTGTTGACCC	AAGTGCAGGCT	CTTGGAGAGC	840
TGTTGGATGG	AGGTGTTAAT	GATGGGGCTG	ATGTGGCGCT	CAATTGACCA	CCCCGGCAAG	900
CTCATTTG	CTCCAGATCT	TGTTCTGGAC	AGGGATGAGG	GGAAATGCGT	AGAAGGAATT	960
CTGGAAATCT	TTGACATGCT	CCTGGCAACT	ACTTCAAGGT	TTCGAGAGTT	AAAACCTCAA	1020
CACAAAGAAT	ATCTCTGTGT	CAAGGCCATG	ATCCTGCTCA	ATTCCAGTAT	GTACCCTCTG	1080
GTCACAGCGA	CCCAGGATGC	TGACAGCAGC	CGGAAGCTGG	CTCACTTGCT	GAACGCCGTG	1140

ACCGATGCTT TGGTTTGGGT GATTGCCAAG AGCGGCATCT CCTCCCAGCA GCAATCCATG 1200
 CGCCTGGCTA ACCTCCTGAT GCTCCTGTCC CACGTCAGGC ATGCGAGGTC TGCCTGA 1257

(2) INFORMATION FOR SEQ ID NO: 21:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 418 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

Met	Asn	Tyr	Ser	Ile	Pro	Ser	Asn	Val	Thr	Asn	Leu	Glu	Gly	Gly	Pro
1										10					15
()															
Gly	Arg	Gln	Thr	Thr	Ser	Pro	Asn	Val	Leu	Trp	Pro	Thr	Pro	Gly	His
															30
()															
Leu	Ser	Pro	Leu	Val	Val	His	Arg	Gln	Leu	Ser	His	Leu	Tyr	Ala	Glu
															45
()															
Pro	Gln	Lys	Ser	Pro	Trp	Cys	Glu	Ala	Arg	Ser	Leu	Glu	His	Thr	Leu
															50
()															
Pro	Val	Asn	Arg	Glu	Thr	Leu	Lys	Arg	Lys	Val	Ser	Gly	Asn	Arg	Cys
															60
()															
Ala	Ser	Pro	Val	Thr	Gly	Pro	Gly	Ser	Lys	Arg	Asp	Ala	His	Phe	Cys
															85
()															
Ala	Val	Cys	Ser	Asp	Tyr	Ala	Ser	Gly	Tyr	His	Tyr	Gly	Val	Trp	Ser
															100
()															
Cys	Glu	Gly	Cys	Lys	Ala	Phe	Phe	Lys	Arg	Ser	Ile	Gln	Gly	His	Asn
															110
()															
Asp	Tyr	Ile	Cys	Pro	Ala	Thr	Asn	Gln	Cys	Thr	Ile	Asp	Lys	Asn	Arg
															130
()															
Arg	Lys	Ser	Cys	Gln	Ala	Cys	Arg	Leu	Arg	Lys	Cys	Tyr	Glu	Val	Gly
															140
()															
Met	Val	Lys	Cys	Gly	Ser	Arg	Arg	Glu	Arg	Cys	Gly	Tyr	Arg	Leu	Val
															165
()															
															175
()															

Arg Arg Gln Arg Ser Ala Asp Glu Gln Leu His Cys Ala Gly Lys Ala
 180 185 190
 Lys Arg Ser Gly Gly His Ala Pro Arg Val Arg Glu Leu Leu Leu Asp
 195 200 205
 Ala Leu Ser Pro Glu Gln Leu Val Leu Thr Leu Leu Glu Ala Glu Pro
 210 215 220
 Pro His Val Leu Ile Ser Arg Pro Ser Ala Pro Phe Thr Glu Ala Ser
 225 230 235 240
 Met Met Met Ser Leu Thr Lys Leu Ala Asp Lys Glu Leu Val His Met
 245 250 255
 Ile Ser Trp Ala Lys Lys Ile Pro Gly Phe Val Glu Leu Ser Leu Phe
 260 265 270
 Asp Gln Val Arg Leu Leu Glu Ser Cys Trp Met Glu Val Leu Met Met
 275 280 285
 Gly Leu Met Trp Arg Ser Ile Asp His Pro Gly Lys Leu Ile Phe Ala
 290 295 300
 Pro Asp Leu Val Leu Asp Arg Asp Glu Gly Lys Cys Val Glu Gly Ile
 305 310 315 320
 Leu Glu Ile Phe Asp Met Leu Leu Ala Thr Thr Ser Arg Phe Arg Glu
 325 330 335
 Leu Lys Leu Gln His Lys Glu Tyr Leu Cys Val Lys Ala Met Ile Leu
 340 345 350
 Leu Asn Ser Ser Met Tyr Pro Leu Val Thr Ala Thr Gln Asp Ala Asp
 355 360 365
 Ser Ser Arg Lys Leu Ala His Leu Leu Asn Ala Val Thr Asp Ala Leu
 370 375 380
 Val Trp Val Ile Ala Lys Ser Gly Ile Ser Ser Gln Gln Gln Ser Met
 385 390 395 400
 Arg Leu Ala Asn Leu Leu Met Leu Leu Ser His Val Arg His Ala Arg
 405 410 415
 Ser Ala

(2) INFORMATION FOR SEQ ID NO: 22:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 34 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

CTTGGATCCA TAGCCCTGCT GTGATGAATT ACAG

34

(2) INFORMATION FOR SEQ ID NO: 23:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 33 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

GATGGATCCT CACCTCAGGG CCAGGCGTCA CTG

33

(2) INFORMATION FOR SEQ ID NO: 24:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1898 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

CACGAATCTT TGAGAACATT ATAATGACCT TTGTGCCTCT TCTTGCAAGG TGTTTCTCA

60

GCTGTTATCT CAAGACATGG ATATAAAAAA CTCACCATCT AGCCTTAATT CTCCTTCCTC

120

CTACAACCTGC AGTCAATCCA TCTTACCCCT GGAGCACGGC TCCATATACA TACCTTCCTC

180

CTATGTAGAC AGCCACCATG AATATCCAGC CATGACATT C TATAGCCCTG CTGTGATGAA	240
TTACAGCATT CCCAGCAATG TCACTAACTT GGAAGGTGGG CCTGGTCGGC AGACCACAAG	300
CCC AAATGTG TTGTGGCAA CACCTGGCA CCTTTCTCCT TTAGTGGTCC ATCGCCAGTT	360
ATCACATCTG TATGCGGAAC CTCAAAAGAG TCCCTGGTGT GAAGCAAGAT CGCTAGAACAA	420
CACCTTACCT GTAAACAGAG AGACACTGAA AAGGAAGGTT AGTGGGAACC GTTGCGCCAG	480
CCCTGTTACT GGTCCAGGTT CAAAGAGGGA TGCTCACTTC TGCGCTGTCT GCAGCGATTA	540
CGCATCGGGA TATCACTATG GAGTCTGGTC GTGTGAAGGA TGTAAGGCCT TTTTTAAAAG	600
AAGCATTCAA GGACATAATG ATTATATTTG TCCAGCTACA AATCAGTGTA CAATCGATAA	660
AAACCGGCCG AAGAGCTGCC AGGCCTGCCG ACTTCGGAAG TGTTACGAAG TGGGAATGGT	720
GAAGTGTGGC TCCCCGGAGAG AGAGATGTGG GTACCGCCTT GTGCGGAGAC AGAGAAGTGC	780
CGACCGAGCAG CTGCACTGTG CCGGCAAGGC CAAGAGAAGT GGCGGCCACG CGCCCCGAGT	840
GCGGGAGCTG CTGCTGGACG CCCTGAGCCC CGAGCAGCTA GTGCTCACCC TCCTGGAGGC	900
TGAGCCGCC C CATGTGCTGA TCAGCCGCC CAGTGC GCCC TTCACCGAGG CCTCCATGAT	960
GATGTCCTG ACCAAGTTGG CCGACAAGGA GTTGGTACAC ATGATCAGCT GGGCCAAGAA	1020
GATTCCCGGC TTTGTGGAGC TCAGCCTGTT CGACCAAGTG CGGCTCTTGG AGAGCTGTTG	1080
GATGGAGGTG TTAATGATGG GGCTGATGTG GCGCTCAATT GACCACCCCG GCAAGCTCAT	1140
CTTGCTCCA GATCTTGTTC TGGACAGGGA TGAGGGAAA TGC GTAGAAG GAATTCTGGA	1200
AATCTTGAC ATGCTCCTGG CAACTACTTC AAGGTTCGA GAGTTAAAAC TCCAACACAA	1260
AGAATATCTC TGTGTCAAGG CCATGATCCT GCTCAATTCC AGTATGTACC CTCTGGTCAC	1320
AGCGACCCAG GATGCTGACA GCAGCCGGAA GCTGGCTCAC TTGCTGAACG CCGTGACCGA	1380
TGCTTGGTT TGGGTGATTG CCAAGAGCGG CATCTCCTCC CAGCAGCAAT CCATGCGCCT	1440
GGCTAACCTC CTGATGCTCC TGTCCCACGT CAGGCATGCG AGTAACAAGG GCATGGAACA	1500
TCTGCTCAAC ATGAAGTGCA AAAATGTGGT CCCAGTGTAT GACCTGCTGC TGGAGATGCT	1560
GAATGCCAC GTGCTTCGCG GGTGCAAGTC CTCCATCACG GGGTCCGAGT GCAGCCCGGC	1620
AGAGGACAGT AAAAGCAAAG AGGGCTCCCA GAACCCACAG TCTCAGTGAC GCCTGGCCCT	1680
GAGGTGAACT GGCCCACAGA GGT CACAAGC TGAAGCGTGA ACTCCAGTGT GTCAGGAGCC	1740

TGGGCTTCAT CTTTCTGCTG TGTGGTCCCT CATTGGTGA TGGCAGGCTT GGTCATGTAC	1800
CATCCTTCCC TCCACCTTCC CAACTCTCAG GAGTCGGTGT GAGGAAGCCA TAGTTCCCT	1860
TGTTAGCAGA GGGACATTG AATCGAGCGT TTCCACAC	1898

(2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 530 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

Met Asp Ile Lys Asn Ser Pro Ser Ser Leu Asn Ser Pro Ser Ser Tyr			
1	5	10	15
Asn Cys Ser Gln Ser Ile Leu Pro Leu Glu His Gly Ser Ile Tyr Ile			
20	25	30	
Pro Ser Ser Tyr Val Asp Ser His His Glu Tyr Pro Ala Met Thr Phe			
35	40	45	
Tyr Ser Pro Ala Val Met Asn Tyr Ser Ile Pro Ser Asn Val Thr Asn			
50	55	60	
Leu Glu Gly Pro Gly Arg Gln Thr Thr Ser Pro Asn Val Leu Trp			
65	70	75	80
Pro Thr Pro Gly His Leu Ser Pro Leu Val Val His Arg Gln Leu Ser			
85	90	95	
His Leu Tyr Ala Glu Pro Gln Lys Ser Pro Trp Cys Glu Ala Arg Ser			
100	105	110	
Leu Glu His Thr Leu Pro Val Asn Arg Glu Thr Leu Lys Arg Lys Val			
115	120	125	
Ser Gly Asn Arg Cys Ala Ser Pro Val Thr Gly Pro Gly Ser Lys Arg			
130	135	140	
Asp Ala His Phe Cys Ala Val Cys Ser Asp Tyr Ala Ser Gly Tyr His			
145	150	155	160

Tyr Gly Val Trp Ser Cys Glu Gly Cys Lys Ala Phe Phe Lys Arg Ser
 165 170 175
 Ile Gln Gly His Asn Asp Tyr Ile Cys Pro Ala Thr Asn Gln Cys Thr
 180 185 190
 Ile Asp Lys Asn Arg Arg Lys Ser Cys Gln Ala Cys Arg Leu Arg Lys
 195 200 205
 Cys Tyr Glu Val Gly Met Val Lys Cys Gly Ser Arg Arg Glu Arg Cys
 210 215 220
 Gly Tyr Arg Leu Val Arg Arg Gln Arg Ser Ala Asp Glu Gln Leu His
 225 230 235 240
 Cys Ala Gly Lys Ala Lys Arg Ser Gly Gly His Ala Pro Arg Val Arg
 245 250 255
 Glu Leu Leu Leu Asp Ala Leu Ser Pro Glu Gln Leu Val Leu Thr Leu
 260 265 270
 Leu Glu Ala Glu Pro Pro His Val Leu Ile Ser Arg Pro Ser Ala Pro
 275 280 285
 Phe Thr Glu Ala Ser Met Met Ser Leu Thr Lys Leu Ala Asp Lys
 290 295 300
 Glu Leu Val His Met Ile Ser Trp Ala Lys Lys Ile Pro Gly Phe Val
 305 310 315 320
 Glu Leu Ser Leu Phe Asp Gln Val Arg Leu Leu Glu Ser Cys Trp Met
 325 330 335
 Glu Val Leu Met Met Gly Leu Met Trp Arg Ser Ile Asp His Pro Gly
 340 345 350
 Lys Leu Ile Phe Ala Pro Asp Leu Val Leu Asp Arg Asp Glu Gly Lys
 355 360 365
 Cys Val Glu Gly Ile Leu Glu Ile Phe Asp Met Leu Leu Ala Thr Thr
 370 375 380
 Ser Arg Phe Arg Glu Leu Lys Leu Gln His Lys Glu Tyr Leu Cys Val
 385 390 395 400
 Lys Ala Met Ile Leu Leu Asn Ser Ser Met Tyr Pro Leu Val Thr Ala
 405 410 415
 Thr Gln Asp Ala Asp Ser Ser Arg Lys Leu Ala His Leu Leu Asn Ala
 420 425 430

Val Thr Asp Ala Leu Val Trp Val Ile Ala Lys Ser Gly Ile Ser Ser
 435 440 445
 Gln Gln Gln Ser Met Arg Leu Ala Asn Leu Leu Met Leu Leu Ser His
 450 455 460
 Val Arg His Ala Ser Asn Lys Gly Met Glu His Leu Leu Asn Met Lys
 465 470 475 480
 Cys Lys Asn Val Val Pro Val Tyr Asp Leu Leu Leu Glu Met Leu Asn
 485 490 495
 Ala His Val Leu Arg Gly Cys Lys Ser Ser Ile Thr Gly Ser Glu Cys
 500 505 510
 Ser Pro Ala Glu Asp Ser Lys Ser Lys Glu Gly Ser Gln Asn Pro Gln
 515 520 525
 Ser Gln
 530

(2) INFORMATION FOR SEQ ID NO: 26:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 30 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

GTGCGGGATCC TCTCAAGACA TGGATATAAA

30

(2) INFORMATION FOR SEQ ID NO: 27:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 25 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

AGTAACAGGG CTGGCGAAC GGTTC

25

(2) INFORMATION FOR SEQ ID NO: 28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

ACTGGCGATG GACCACTAAA GG

22